

Claims

What is claimed is:

5

1. A text to XML transformer, comprising:

a transformer program having a plurality of executable statements; and
a processor for executing the transformer program and converting an

10

input text document into an XML document wherein the XML document does not
contain every element that was in the input text.

2. The transformer of claim 1, wherein the text document is a
structured text document.

15

3. The transformer of claim 1, wherein the text document is a semi-
structured text document.

4. The transformer of claim 1, wherein the input text document has at
least two formats.

20

5. The transformer of claim 4, wherein the text to XML commands
include a field separator command that defines a field separator in the text
document.

25

6. The transformer of claim 5, wherein the field separator is a comma.

7. The transformer of claim 5, wherein the field separator is a regular
expression.

8. The transformer of claim 4, wherein the text to XML commands include a match command that requires a field in the input text document to match a character string or a record is skipped.

5

9. The transformer of claim 4, wherein the text to XML commands include a tree hierarchy command.

10. The transformer of claim 1, wherein the input text document is a streaming text.

10

11. The transformer of claim 1, wherein the XML document is a streaming XML.

12. The transformer of claim 1, further including a wizard that has a number of queries that are used to defined the transformer program.

15

13. The transformer of claim 1, wherein the input text document is from a legacy system and an output is to an XML system.

20

14. A process for converting text to XML, comprising the steps of:

a) defining a transformer program having a plurality of executable statements, wherein one of the plurality of executable statements contains a command that matches a regular expression and takes an action;

25

b) receiving a text stream;

c) executing the transformer program to convert the text stream into an XML stream.

15. The process of claim 14, wherein the step (a) further includes the step of:

5 a1) selecting a text to XML wizard.

16. The process of claim 14, wherein step (a) further includes the steps of:

10 a1) selecting a field separator command that defines a field separator in the text stream.

17. The process of claim 16, wherein step (a1) further includes the steps of:

15 i) defining the field separator as a regular expression.

18. The process of claim 14, wherein step (b) further includes the steps of:

20 b1) receiving the text stream having two or more formats.

19. A text to XML transformer, comprising:

a wizard for creating a transformer document;

5 the transformer document having a plurality of statements formed by a text to XML computer language; and

a processor for executing the transformer document and converting an input text document into an XML document.

10 20. The transformer of claim 19, wherein the text to XML computer language includes a section command to define a section.

21. The transformer of claim 21, wherein the section command uses a regular expression match to define the section.

15

20